

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using SW model

Run on: August 21, 2002, 15:44:28 ; Search time 13.43 seconds  
(without alignments)  
986,008 Million cell updates/sec

Title: US-09-780-576-2

Sequence: 1 MOAVDNLTSAPCNTSLCTRP . . . . . SUNKKKKBUUGUPNPTPM 342

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

al number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSPROT\_40:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	772	43.4	338	1 K101_HUMAN	Q15391 homo sapien
2	694	39.0	305	1 K101_RAT	Q15881 rattus norv
3	488	27.4	319	1 H961_HUMAN	Q14626 homo sapien
4	411.5	23.1	375	1 GPR1_MOUSE	Q9116 mus musculu
5	409.5	23.0	381	1 GPR1_HUMAN	Q9116 mus musculu
6	394	22.2	342	1 PAFR_HUMAN	P25105 homo sapien
7	392.5	22.1	342	1 PAFR_CANPO	P21556 cavia porce
8	379.5	21.3	361	1 EB12_HUMAN	P32249 homo sapien
9	367	20.6	345	1 CLT2_PIG	Q95003 sus scrofa
10	366	20.6	341	1 PAFR_MOUSE	Q62035 mus musculu
11	362	20.4	341	1 PAFR_RAT	P46002 rattus norv
12	357.5	19.5	308	1 P2Y5_CHICK	P32250 gallus gall
13	347.5	19.5	346	1 P2Y8_XENIA	Q95075 homo sapien
14	347.5	19.5	537	1 P2Y8_HUMAN	P79928 xenopus lae
15	344	19.3	309	1 CLT2_MOUSE	Q92041 mus musculu
16	339.5	19.1	359	1 AG2R_BOVIN	Q92459 rattus norv
17	339	19.0	344	1 P2Y5_HUMAN	P43657 homo sapien
18	338	18.8	359	1 AG2R_SHEEP	Q77580 vis aries
19	334	18.5	339	1 GPRH_HUMAN	Q13304 homo sapien
20	329.5	18.5	359	1 AG2S_HUMAN	Q13725 homo sapien
21	329	18.5	340	1 AG2S_MOUSE	Q95002 sus scrofa
22	328	18.4	359	1 AG2R_PIG	P30555 sus scrofa
23	327	18.4	359	1 AG2R_CANFA	P43240 canis fami
24	326	18.3	359	1 PAR2_MOUSE	P55086 mus musculu
25	325.5	18.3	359	1 PAR2_MOUSE	P55086 mus musculu
26	325	18.3	359	1 AG2R_RABIT	P40556 homo sapien
27	321	18.1	359	1 AG2R_HUMAN	Q9116 mus musculu
28	320	18.0	370	1 P2Y9_CANPO	Q9116 mus musculu
29	320	18.0	370	1 P2Y9_HUMAN	Q15743 homo sapien
30	319.5	18.0	365	1 GPR6_HUMAN	P25079 rattus norv
31	319	17.9	359	1 AG2R_RAT	P29088 rattus norv
32	319	17.9	359	1 AG2S_RAT	P29754 mus musculu
33	318	17.9	359	1 AG2R_MOUSE	P29754 mus musculu

RESULT	1	STANDARD	PRT	338 AA
1	K101_HUMAN			
2	K101_HUMAN			
3	K101_HUMAN			
4	K101_HUMAN			
5	K101_HUMAN			
6	K101_HUMAN			
7	K101_HUMAN			
8	K101_HUMAN			
9	K101_HUMAN			
10	K101_HUMAN			
11	K101_HUMAN			
12	K101_HUMAN			
13	K101_HUMAN			
14	K101_HUMAN			
15	K101_HUMAN			
16	K101_HUMAN			
17	K101_HUMAN			
18	K101_HUMAN			
19	K101_HUMAN			
20	K101_HUMAN			
21	K101_HUMAN			
22	K101_HUMAN			
23	K101_HUMAN			
24	K101_HUMAN			
25	K101_HUMAN			
26	K101_HUMAN			
27	K101_HUMAN			
28	K101_HUMAN			
29	K101_HUMAN			
30	K101_HUMAN			
31	K101_HUMAN			
32	K101_HUMAN			
33	K101_HUMAN			
34	K101_HUMAN			
35	K101_HUMAN			
36	K101_HUMAN			
37	K101_HUMAN			
38	K101_HUMAN			
39	K101_HUMAN			
40	K101_HUMAN			
41	K101_HUMAN			
42	K101_HUMAN			
43	K101_HUMAN			
44	K101_HUMAN			
45	K101_HUMAN			

## ALIGNMENTS

RESULT	1	STANDARD	PRT	338 AA
1	K101_HUMAN			
2	K101_HUMAN			
3	K101_HUMAN			
4	K101_HUMAN			
5	K101_HUMAN			
6	K101_HUMAN			
7	K101_HUMAN			
8	K101_HUMAN			
9	K101_HUMAN			
10	K101_HUMAN			
11	K101_HUMAN			
12	K101_HUMAN			
13	K101_HUMAN			
14	K101_HUMAN			
15	K101_HUMAN			
16	K101_HUMAN			
17	K101_HUMAN			
18	K101_HUMAN			
19	K101_HUMAN			
20	K101_HUMAN			
21	K101_HUMAN			
22	K101_HUMAN			
23	K101_HUMAN			
24	K101_HUMAN			
25	K101_HUMAN			
26	K101_HUMAN			
27	K101_HUMAN			
28	K101_HUMAN			
29	K101_HUMAN			
30	K101_HUMAN			
31	K101_HUMAN			
32	K101_HUMAN			
33	K101_HUMAN			
34	K101_HUMAN			
35	K101_HUMAN			
36	K101_HUMAN			
37	K101_HUMAN			
38	K101_HUMAN			
39	K101_HUMAN			
40	K101_HUMAN			
41	K101_HUMAN			
42	K101_HUMAN			
43	K101_HUMAN			
44	K101_HUMAN			
45	K101_HUMAN			

P33535 rattus norv  
P29755 mus musculu  
Q91163 rattus norv  
Q63645 rattus norv  
P55414 homo sapien  
Q95247 sus scrofa  
Q15210 meriones un  
P79350 bos taurus  
Q08675 mus musculu  
P32303 xenopus lae  
P35373 xenopus lae  
Q00254 homo sapien

PROBABLE ORTHOLOG OF RAT VTR 15-20.  
-1- FUNCTION: ORPHAN RECEPTOR.  
-1- SUBCELLULAR LOCATION: Integral membrane protein.  
-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: D13626; BAA02791.1; -  
CC GCRdb; GCR\_0594; -  
CC InterPro: IPR000276; GPCR\_Rhodpsn.  
CC Pfam: PF00001; 7tm\_1; 1.  
CC PROSITE: PS00237; G-PROTEIN\_RECPT\_F1\_1; FALSE\_NEG.  
CC PROSITE: PS00262; G-PROTEIN\_RECPT\_F1\_2; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein.

CC DOMAIN 1  
CC 1 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 2 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 3 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 4 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 5 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 6 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 7 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 8 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 9 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 10 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 11 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 12 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 13 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 14 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 15 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 16 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 17 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 18 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 19 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 20 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 21 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 22 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 23 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 24 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 25 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 26 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 27 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 28 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 29 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 30 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 31 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 32 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 33 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 34 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 35 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 36 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 37 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 38 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 39 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 40 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 41 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 42 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 43 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 44 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)  
CC 45 (POTENTIAL)  
CC CYTOPLASMIC (POTENTIAL)

FT DOMAIN 256 278 EXTRACELLULAR (POTENTIAL)  
 FT TRANSMEM 279 299 7 (POTENTIAL)  
 FT DOMAIN 300 338 CYTOPLASMIC (POTENTIAL)  
 FT CARBOHYD 3 3 N-LINKED (GLCNAC) (POTENTIAL)  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC) (POTENTIAL)  
 SQ SEQUENCE 338 AA; 38971 MW; 8DBE7C782CB4755AD CRC64;

Query Match 43 44: Score 772; DB 1; Length 338;  
 Best Local Similarity 47.5%; Pred. No. 3,4e-45;  
 Matches 149; Conservative 60; Mismatches 105; Indels 0; Gaps 0;

DB 3 NSTSTQPPDESCSONLITTOIIPVLCMTFAGILNGVSMFFVYPSKSFITLKN 62  
 QY 66 TVISDLMLTLPFKRLSDAKLGTGPTFVQVTSVFETFWYISISFLGITTIDRYK 125  
 DB 63 IVADFVMSLTFFPKRLSDGSLGPMQNVFVGSVAVIFVNMVYSIAFPGLISPDHYK 122  
 QY 126 TTRPKTSNPNKLLGAKILSVVIAFMFLSLPMMLITNRPDRKNKCKSEFLSEGLV 185  
 DB 123 IVKPLMTSFIQSVSKSLSVIWMMLMLAVPNLLITNDSVPEVYIKLELSLGRK 182  
 QY 186 WHEIVNYICQVIFWIFINFLIVCYTLITKELRYSTVRIKGVKVKVKKVPIIAVE 245  
 DB 183 WHKASNTYFAVIFWIFELLIVFYATIKKIFKSHLSKSNISVKKSSRNIFSEIVEF 242  
 QY 246 FICFPEFHAFIPTLSLTREVFDCTAENTLFYKESLMLTSLNACLDPIYFLKSF 305  
 DB 243 FVCEPFIHAFIPTKSTQTEAHYSQSKELRYKMEPTLLLSAANVCLPIYFLCOPE 302  
 QY 306 NSLISMLKCPNSA 319  
 DB 303 REILCKKHIFPLKA 316

## RESULT 2

ID K101\_RAT STANDARD; PRT; 305 AA.  
 AC 035881;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Probable G protein-coupled receptor VTR 15-20.  
 OS Rattus norvegicus (Rat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 CC NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.  
 RA Chaiton M.E., Williams A.S., Fogliano M., Sweetnam P.M., Duman P.S.;  
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: ORPHAN RECEPTOR.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC PROBABLE ORTHOLOG OF HUMAN KIAA0001.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: U76206; AAB71745.1;  
 CC GCRDB: GCR\_2520;  
 CC InterPro: IPR000276; GPCR\_Rhodpsn.  
 CC Pfam: PF00001, 7tm\_1, 1.  
 CC PROSITE: PS00237; G\_PROTEIN\_RECP\_FL\_1; FALSE\_NEG.  
 CC PROSITE: PS50262; G\_PROTEIN\_RECP\_FL\_2; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 29 EXTRACELLULAR (POTENTIAL)  
 FT TRANSMEM 30 50 1 (POTENTIAL)  
 FT DOMAIN 51 55 CYTOPLASMIC (POTENTIAL)  
 FT TRANSMEM 56 76 2 (POTENTIAL)  
 FT DOMAIN 77 96 EXTRACELLULAR (POTENTIAL)  
 FT TRANSMEM 97 117 3 (POTENTIAL)  
 FT DOMAIN 118 139 CYTOPLASMIC (POTENTIAL)  
 FT TRANSMEM 140 160 4 (POTENTIAL)  
 FT DOMAIN 161 188 EXTRACELLULAR (POTENTIAL)  
 FT TRANSMEM 189 209 5 (POTENTIAL)  
 FT DOMAIN 210 234 CYTOPLASMIC (POTENTIAL)  
 FT TRANSMEM 235 255 6 (POTENTIAL)  
 FT DOMAIN 256 278 EXTRACELLULAR (POTENTIAL)  
 FT TRANSMEM 279 299 7 (POTENTIAL)  
 FT CARBOHYD 300 305 CYTOPLASMIC (POTENTIAL)  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC) (POTENTIAL)  
 SQ SEQUENCE 305 AA; 34800 MW; 640546A68ECC7A3 CRC64;

Query Match 39 0%; Score 694; DB 1; Length 305;  
 Best Local Similarity 45.2%; Pred. No. 5.1e-40;  
 Matches 132; Conservative 61; Mismatches 99; Indels 0; Gaps 0;

DB 4 VDMLSAPGNTSLCTRDYKTOVLPFLVLPFVGLITNGLAMRIFFQIRKSNFIITL 63  
 DB 1 MONTTTPPKQKNTLITTOIIPMLYQVEINSLNGISGMFFVPSKSFITL 60  
 QY 64 KNVVISDLMLTLPFKRLSDAKLGTGPTFVQVTSVFETFWYISISFLGITTIDRY 123  
 DB 61 KNIVADFMGLTFFPKRLSDGSLGPMQNVFVGSVAVIFVNMVYSIAFPGLISPDY 120  
 QY 124 QKTRPKTSNPNKLLGAKILSVVIAFMFLSLPMMLITNRPDRKNKCKSEFLSEGLV 183  
 DB 121 YKIVKPLWSISVSVKSVLWVLMMLAVENIILITNSVKKVITMLQMELEKNEG 180  
 QY 184 LVWHEIVNYICQVIFWIFINFLIVCYTLITKELRYSTVRIKGVKVKVKKVPIIAVE 243  
 DB 181 RKWHKASNTYFAVIFWIFELLIVFYATIKKIFKSHLSKSNISVKKSSRNIFSEIVEF 240  
 QY 244 VFCEPFIHAFIPTKSTQTEAHYSQSKELRYKMEPTLLLSAANVCLPIYFLCOPE 295  
 DB 241 AVACAPAHVAFIPTKSTQTEAHYSQSKELRYKMEPTLLLSAANVCLPIYFLCOPE 292

## RESULT 3

ID H963\_HUMAN STANDARD; PRT; 319 AA.  
 AC 014626;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable G protein-coupled receptor H963.  
 OS Homo sapiens (human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 CC NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
 RA TISSUE-Peritoneal blood monocytes;  
 RX MEDLINE=9803606; PubMed=9370294;  
 RA Jacobs K.A., Collins-Racie L.A., Colbert M., Duckett M.,  
 RA Golden-Fleet M., Reller K., Kriz R., LaValle E.R., Metberg D.,  
 RA Spaulding V., Stovek J., Williamson M.J., McCoy J.M.;  
 RT "A genetic selection for isolating cDNAs encoding secreted proteins";  
 RL Gene 198:289-296(1997).

CC -!- FUNCTION: ORPHAN RECEPTOR.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -